



```

Seq primer: -40m13 fwd. ET from AmerSham
High quality sequence stop: 1.
Location/Qualifiers
1. .106
/oranism="Homo sapiens"
/db_xref="taxon:9606"
/cclone="IMAGE:1699097"
/cclone_lib="Gessler Wilms tumor"
/sex="pooled (6)"
/lab_host="DH10B"
/clone="Vector: pSP6K11; Site_1: SalI; Site_2: NotI; RNA
was prepared from a pool of 6 anonymous Wilms' tumor RNAs.
RNA was prepared by acid-phenol, followed by one round of
oligo dT selection. cDNA library preparation was with
the BRL/Life Tech. Superscript plasmid system. An
oligo-dT NotI primer for first strand synthesis generated
qacacccccc(11n at the 3' end of the clones. A 5' SalI
adaptor was used with sequence 5'-gtcaccacacgtcc-3'.
Resulting cDNAs were size selected (average size 2 kb),
NotI digested, and ligated into NotI/SalI-cut pSP6K11.
Library was constructed by Dr. Manfred Gessler."
22 a 26 c 28 g 30 t

```

```

/db_xref="taxon:9606"
/clone="IMAGE:169097"
/clone_lib="Gessler Wllms tumor"
/sex="pooled (6)"
/lab_host="DH10B"
/notes="Vector: pSmaK11; Site 1: SalI; Site 2: NotI; RNA
was prepared from a pool of 6 anaplastic Wllms' tumor RNAs.
RNA was prepared by acid-phenol, followed by one round of
oligo dT selection. cDNA library preparation was with
the BRL/Life Tech. Superscript plasmid system. An
oligo-dT NotI primer for first strand synthesis generated
gaagacgcccttga at the 3' end of the clones. A 5' SalI
adaptor was used with sequence 5'-gtcagcagcagctcc-3'.
Resulting cDNAs were size selected (average size 2 kb),
NotI digested, and ligated into NotI/SalI-cut pSmaK11.
Library was constructed by Dr. Manfred Gessler."
22 a 26 c 28 g 30 i

```

```

BASE COUNT      22 a      26 c      30 l
ORIGIN

Alignment Scores:
      Pred. NO.:      1,34e+03      length:      106
      Score:      24.00      Matches:      5
      Percent Similarity:      100.00%      Conservative:      0
      Percent Local Similarity:      100.00%      Mismatches:      0
      Query Match:      100.00%      Indels:      0
      PH:      9      Gaps:      0

US-09-856-070-18 (1-5) x AI003501 (1-106)

```

US-09-856-070-18 (1-5) x A1003501 (1-106)	
2Y	1 LysGluGluLeuMet. 5 
Db	64 AAGCAAGCACTGATG 50 
RESULT 3	
AZ825768	
LOCUS	
DEFINITION	
AZ825768	100 bp
linear	SSS 20-FEB-2001
clone UGCG2M01G08 F, DNA sequence.	
library Mus musculus genomic	

clone UGCGC2M0101G08 F, DNA sequence.  
A7825768  
A7825768.1 GI:12995676  
GSS.  
house mouse.  
Mus musculus  
ORGANISM  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 109)  
AUTHORS  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly,  
M., Rose,M., Rose,R., Stokes,K., Tingey,A., von Niederhausern,A.  
and Wright,D. Weiss,R.  
TITLE  
Mouse whole genome scaffolding with paired end reads from 10kb

**TITLE**  
 Mouse whole genome scaffolding with paired end reads from 10kb  
 plasmid inserts  
**JOURNAL**  
 Republished (2000)  
**COMMENT**  
 Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., Salt, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: d3em@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0101 Row: C Column: 08  
 Seq primer: GTTTGTAAGGACAGGAGT  
 Class: plasmid ends  
 High quality sequence stop: 109.  
 Location/Qualifiers  
 1..109

source 1..109

/organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="imgc2w0101004"  
 /clone\_lib="Mouse 10kb plasmid UUCGIM library"  
 /sex="Male"  
 /lab\_host "E. Coli strain XL10-Gold, Tl-resistant, P-"  
 /note="Vector: pW042nv, Purified genomic DNA from M.  
 musculus C57BL/6J (male) was obtained from the Jackson  
 Laboratory Mouse DNA resource  
 (http://www.jax.org/resources/documents/dnares/) The DNA  
 was hydrodynamically sheared by repeated passage through a  
 0.005 inch orifice at constant velocity. The sheared DNA  
 was blunt end-repaired with T4 DNA polymerase and T4  
 polynucleotide kinase. Adaptor oligonucleotides were  
 ligated to the blunt ends in high molar excess. The  
 adaptor DNA was purified and size-selected for a 9.5 to  
 10.5 kb range using preparative agarose gel  
 electrophoresis. Vector DNA was prepared from a derivative  
 of pW042 (g114732114[qlb]AF129072.1), a copy-number  
 inducible derivative of plasmid p1. The vector was ligated  
 with adaptors complementary to the insert adaptors and  
 purified. The sheared, adaptor mouse DNA was annealed to  
 adaptor vector DNA, and transformed into  
 chemically competent E. coli XL10-Gold (Stratagene) cells  
 and selected for ampicillin resistance."

BASE COUNT 29 a 15 c 34 q 31 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.4e+03 Length: 109  
 Score: 24.00 Matches: 5  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 17 Gaps: 0

US-09-856-070-18 (1-5) x A2425768 (1-109)

QY 1 LysGluGluLeuMet 5  
 LOCUS  
 |||||||  
 DB 79 AAGGAAGATTCATG 93

RESULT 4  
 AW946335 115 bp mRNA linear EST 31-MAY-2000  
 DEFINITION KC2-ET0018-206400-012-007 ET0018 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION AW946335  
 VERSION AW946335.1 GI:8124101  
 KEYWORDS EST.  
 SOURCE human.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catartini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 115)  
 AUTHORS Dias Neto, E., Garcia Correa, R., Vetrjovsky-Almeida, S., Arlones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Paiva, S.S., Simpson, D.H.,  
 Brunstein, A., de Oliveira, P.S., Bucher, P., Tomazetti, V., O'Hare  
 M.J., Soares, F., Arentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags

TITLE  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 MEDLINE 20202663  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01504-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/11CR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml.pl?1-512-002-ET0018 290  
 400-012-007/856-090-04-29a14-1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 9  
 High quality sequence stop: 115.  
 Location/Qualifiers

FEATURES  
 source

1..115  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="p0018"  
 /dev\_stage="Adult"  
 /note="Origin: lung\_tumor. Vector: puc18. Site: 1: Small;  
 Site: 2: Small; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

BASE COUNT 38 a 20 c 24 q 33 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.51e+03 Length: 115  
 Score: 24.00 Matches: 5  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 16 Gaps: 0

US-09-856-070-18 (1-5) x AW946335 (1-115)

QY 1 LysGluGluLeuMet 5  
 LOCUS  
 |||||||  
 DB 10 AAGGAAGATTCATG 24

RESULT 5  
 AZ761701/c

LOCUS AZ761701 127 bp DNA linear GSS 16-FEB-2001  
 DEFINITION IM0556707F Mouse 10kb plasmid UUCGIM library Mus musculus genomic  
 clone UUCGIM0556J07 F, DNA sequence.

ACCESSION AZ761701  
 VERSION AZ761701.1 GI:12870903

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 127)  
 AUTHORS Dunn, D., Anyagi, A., Barber, M., Beasorn, T., Duval, B., Hamil, C.,  
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly  
 M., Rose, M., Rose, R., Stokes, R., Tinney, A., von Niederhausern, A.  
 and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

UNPUBLISHED (2000)

CONTACT: Robert B. Weiss

UNIVERSITY OF UTAH

RM. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

TEL: 801 585 5606

FAX: 801 585 7177

EMAIL: dunn@genetics.utah.edu

INSERT LENGTH: 10000 STD ERROR: 0.00

PLATE: 0556 ROW: J COLUMN: 07

SEQ PRIMER: GCTTCAAAACGACGCCACAT

CLASS: plasmid ends

HIGH QUALITY SEQUENCE STOP: 127.

LOCATION/QUALIFIERS

1..127

/organism="Mus musculus"

/strain="c57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="MUGGIM0556J07"  
 /clone\_lib="Mouse 10kb plasmid MUGGIM library"  
 /sex="Male"  
 /lab\_host="E. coli strain XL10-Gold, T1-resistant, F-"  
 /note="Vector: pWD42nv; Purified genomic DNA from M  
 muscululus c57BL/6J (male) was obtained from the Jackson  
 Laboratory Mouse DNA Resource  
 (http://www.jax.org/resources/documents/dnates/). The DNA  
 was hydrolytically sheared by repeated passage through a  
 0.005 inch orifice at constant velocity. The sheared DNA  
 was blunt end-repaired with T4 DNA polymerase and T4  
 polynucleotide kinase. Adaptor oligonucleotides were  
 ligated to the blunt ends in high molar excess. The  
 adaptor DNA was purified and size-selected for a 9.5 to  
 10.5 kb range using preparative agarose gel  
 electrophoresis. Vector DNA was prepared from a derivative  
 of pWD42 (414732114[ab]AF12672.1), a copy number  
 inducible derivative of plasmid p1. The vector was ligated  
 with adaptors complementary to the insert adaptors and  
 purified. The sheared, adaptor mouse DNA was annealed to  
 adaptor vector DNA, and transformed into  
 chemically-competent E. coli XL10-Gold (Stratagene) cells  
 and selected for ampicillin resistance."

BASE COUNT 24 a 33 c 28 3 42 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.74e+03 Length: 127  
 Score: 24.00 Matches: 5  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 17 Gaps: 0

US-09-856-070-18 (1-5) x AZ761701 (1-127)

QY 1 LysGluGluLeuMet 5

DB 85 AAGGAGCAATGATG 71

RESULT 6

BM452857

LOCUS BM452857 130 bp mRNA linear EST 07-JAN-2002  
 DEFINITION 168604.Y1 HR85 islet Homo sapiens cDNA 5' similar to SW:CBPH\_HUMAN  
 176870 CARBOXYPEPTIDASE H PRECURSOR 1, mRNA sequence.

ACCESSION BM452857

VERSION BM452857.1 GI:18085215

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 130)

AUTHORS Melton D., Brown J., Kenly G., Permutt, A., Lee C., Kaestner, K.,  
 Hillier, J., Marra, M., Papadopoulos, J., Gradwohl, G., Clifton, S.,  
 Lemishka, I., Searles, M., Wroblewski, J., Wylic, P., Martin, J., Hlistain, A.,  
 Schmitt, A., Theising, H., Ritter, E., Renko, I., Bennett, J., Cardenas  
 M., Gibbons, M., McCann, R., Cole, R., Tsadareishvili, R., Williams, T.,  
 Jackson, Y., and Bowers, Y.

Endocrine Pancreas Consortium

Unpublished (2000)

Other ESTs: 168604.x1

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel.: 617-495-1812

Fax: 617-495-8557

Email: dmelton@chp.harvard.edu

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by.

Washington University Genome Sequencing Center For information on  
 obtaining a clone please contact: Dr. Hiroshi Inoue  
 (hinoue@im.wustl.edu)

Trace considered overall poor quality

Seq primer: -40RP from Gibco

High quality sequence stop: 1.

Location/Qualifiers

source

1..130

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="HR85 islet"

/tissue\_type="Purified pancreatic islet"

/lab\_host="DH10B"

/note="Clonab: Pancreas; Vector: pBluescript SK(-); Site\_1:  
 NotI; Site\_2: XhoI; cDNA made by oligo-dT priming. -kb. 5'

Size: selected on agarose gel. Average insert size: -kb. 5'

XhoI site was destroyed after directional cloning.

Amplified once. Contact information: Hiroshi Inoue, MD,

Medical-115m Flv (Alan Feinberg Lab), Washington University

School of Medicine, Box 8127, 660 South Euclid Ave., St.

Louis, MO 63110. E mail: hinoue@im.wustl.edu. Tel:

314-362-1916. Fax: 314-747 2692."

BASE COUNT 43 a 14 c 37 g 36 t

ORIGIN

Alignment Scores:

Pred. No.: 1.8e+03 Length: 130

Score: 24.00 Matches: 5

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 13 Gaps: 0

US-09-856-070-18 (1-5) x HM352857 (1-130)

QY 1 LysGluGluLeuMet 5

DB 80 AAGAGCAATGATG 94

RESULT 7

BM364004

LOCUS BM364004

DEFINITION 161-570278-200009 110-d11\_2 ST0278 Homo sapiens cDNA, mRNA

sequence.

ACCESSION BM364004

VERSION BM364004.1 GI:21036015

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 130)

AUTHORS Dias Neto, E., Garcia Correa, K., Verjovsky-Almeida, S., Briones, M.R.,

Naqai, M.A., da Silva, W. Jr., Zaqui, M.A., Bordin, S., Costa, F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,

Brunstein, A., deoliveira, P.S., Bucher, P., Jonqueel, C.V., O'Hare

M.J., Soares, P., Brentani, R.R., Reis, L.F., de Souza, S.J., and

Simpson, A.J.

Shotgun sequencing of the human transcriptome with over expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

CONTACT: Simpson A.J.G.

Laboratory of Cancer Genetics

Madhy Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel.: 55-11-2764922

Fax: 55-11-2707001

Email: asimpson@mdwq.org.br

This sequence was derived from the FAPESP/ICP Human Cancer Genome

Project. This entry can be seen in the following URL:

(http://www.bicm.wustl.edu/scripts/gethtml2.pl?1 603-ST0278-

200000-110-411-581-4-0000-00-20614-1)

Seq primer: puc 18 forward  
High quality sequence stop: 130.

FEATURES  
Location/Qualifiers

1..130  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="ST0278"  
/dev\_stage="Adult"  
Site 2: Small: A mini library was made by cloning products derived from ORESIES PCR (U.S. letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) tissues into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions.  
44 a 25 c 35 g 26 t

BASE COUNT  
ORIGIN

Alignment Scores:  
Prod. No.: 1.8e+03 Length: 130  
Score: 24.00 Matches: 5  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-09-856-070-18 (1-5) x HQ364004 (1-130)

QY 1 LysGluGluLeuMet 5  
|||||  
DB 103 AAGAGAGAGCTGATG 117

RESULT 8

W86729 136 bp mRNA linear EST 02-FEB-1997  
LOCUS  
DEFINITION  
clone IMAGE 416756 5' similar to SW1008\_MOUSE P24556 BRAIN PROTEIN  
DN38 ; mRNA sequence.

ACCESSION W86729  
VERSION W86729.1 GI:1400477  
KEYWORDS EST.  
SOURCE human.

ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 136)  
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, R.,  
Chissoe, S., Dietrich, N., DuBuque, T., Favell, A., Gish, W., Hawkins,  
M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore,  
R., Morris, M., Parsons, J., Prange, C., Rifkin, J., Rohlfing, T.,  
Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E.,  
Underwood, K., Weidmann, P., Waterston, R., Wilson, R. and Maria, M.  
Generation and analysis of 280,000 human expressed sequence tags  
Genome Res 6 (5), 807-828 (1996)  
9704478

TITLE

JOURNAL  
MEDLINE  
COMMENT  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu

This clone is available royalty-free through EMBL; contact the  
IMAGE Consortium (infoimage.llnl.gov) for further information.  
Trace considered overall poor quality  
Possible reversed clone similarity on wrong strand  
Insert Length: 957 5' Error: 0.00  
Seq primer: mb.PE3A+ET  
High quality sequence stop: 1.

FEATURES  
Location/Qualifiers

1..136  
/organism="Homo sapiens"  
/db\_xref="GDB:132526"

/db\_xref="taxon:9606"  
/clone\_lib="Soares\_fetal\_liver\_spleen\_INFUS\_S1"  
/sex="male"  
/dev\_stage="20 week post conception fetus"  
/lab\_host="EM103 (ampicillin resistant)"  
/note="organ: Liver and Spleen; Vector: pT73D (Pharmacia)  
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;  
this is a subtracted version of the original Soares fetal  
liver spleen INFUS library. 1st strand cDNA was primed  
with a Pac I oligo(dT) primer 15'  
AATGGAGAGATTAATTAAAGATCTTTTTTTTTTTTTTTT 3',  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified pT73 vector. Library  
went through one round of normalization. Library  
constructed by Rento Soares and M.Fatima Ronaldo."

BASE COUNT 32 a 25 c 22 g 54 t 3 others

Alignment Scores:  
Prod. No.: 1.92e+03 Length: 136  
Score: 24.00 Matches: 5  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-09-856-070-18 (1-5) x W86729 (1-136)

QY 1 LysGluGluLeuMet 5  
|||||  
DB 96 AAGAGAGATTAATG 82

RESULT 9

W8774194 137 bp mRNA linear EST 20-SEP-2000  
LOCUS  
DEFINITION  
X91-00000-180000-14-b10 OM3009 Homo sapiens cDNA, mRNA sequence.  
W8774194  
VERSION W8774194.1 GI:10227758  
KEYWORDS EST.  
SOURCE human.

ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 137)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,  
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.

Shotgun sequencing of the human transcriptome with 38K expressed  
sequence tags  
Proc Natl Acad Sci U S A 97 (7): 3491-3496 (2000)  
20202663  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Poa Prof. Antonio Prudente 109, 4 andar, 01500-010, Sao Paulo SP,  
Brazil

Tel: +55-11-3704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/ILIC Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?lib=at2-MPL-000094-180  
500-014-1000-05-18&t4-1)  
Seq primer: puc 18 forward  
High quality sequence stop: 40  
High quality sequence stop: 137.  
Location/Qualifiers

1..137

FEATURES  
source

```

Alignment Scores:
Pred. No.: 1.96e+03      Length: 148
Score: 24.00      Matches: 5
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%      Indels: 0
DH: 17      Gaps: 0
D5-09-066-070-10 (1-5) x B8H11366 (1-138)

```

```

Qy      1  lysGlnGluLeuMet 5
|||||
Db      67  AAGCAGCAGCGTCGATG 53

RESULT 11
AW946443
LOCUS      AW946443
DEFINITION  EC2:PT0019-090509-012 f12 PT0019
ACCESSION  AW946443
VERSION    AW946443.1 GI:1124212
KEYWORDS   EST.
SOURCE     human

```

ORGANISM	Homo Sapiens
REFERENCE	Fukuyama; Motazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea (bases 1 to 139)
AUTHORS	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,N.A., da Silva,W. Jr., Zazo,M.A., Boidin-Goldman,G.H., Carvalho,A.F., Matsumura,A., Brunstein,A., doOliveira,P.S., Hucher,P., Jones,M.J., Soares,F., Brentani,R.R., Reis,L.F., deSimpson,A.J.
TITLE	Shotgun sequencing of the human transcriptome sequence tags
JOURNAL:	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496
MEDLINE:	200002643

20202003  
 CONTACT: SIMPSON A.J.G.  
 LABORATORY OF CANCER GENETICS  
 LUDWIG INSTITUTE FOR CANCER RESEARCH  
 Rua Prof. Antonio Prudente 109, 4 andar, 01500-000  
 Brazil  
 Tel: +55-11-3704922  
 Fax: +55-11-3707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the EAPESP/LICRF  
 Project. This entry can be seen in the following  
 (<http://www.ludwig.org.br/scripts/gethtml2.pl>)  
 509-012-12653-2000-05-08&t4-1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 13  
 High quality sequence stop: 139  
 Location/Qualifiers  
 1..139  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="ET0019"  
 /seq\_id="34844"  
 FEATURES  
 SOURCE

```

Zoev-Shagel - Adult
Zoev-Grigoriy - Adult
Site_2: Smal; A mini-library was made
derived from ORESTES PCR (U.S. Letter
No. 196,716 - Ludwig Institute for Cancer
Research) into the pUC 18 vector. Reveal
tissue mRNA and cDNA amplification were
low stringency conditions.*
48 a      24 c      31 g      36 t
ORIGIN
Alignment Scores:
Score..          1.98e+03      Length:          139
Inserted No.:    24.00        Matches:         5
Percent Similarity: 100.00%     Conservative:    0
Best Local Similarity: 100.00%   Mismatches:     0

```

```

Query Match: 100.00%      Indels: 0
DB:          10           Gaps: 0

US-09-856-070-18 (1-5) x AW946443 (1-149)

QY 1 LysGluGluLeuMet 5
DB 8 AAGGAGAGCTGATG 22

RESULT 12
LOCUS BH101135
DEFINITION BPCL-24-333H1.TVR BPCL 24 Mus musculus genomic clone BPCL 24 333H1,
DNA sequence.
ACCESSION BH101135
VERSION BH101135.1 GI:14925696
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae; Mus
1 (bases 1 to 143)
AUTHORS Zhao, S., Niemman, W., Malek, J., Shatsman, S., Akinret, P., Levins, M.,
Tsogayev, G., Giet, K., Krol, M., Shwartsbeyn, A., Gebreyes, E.,
Russell, D., de Jong, P. and Fraser, C.M.
TITLE Mouse HAC End Sequences from Library BPCL-24
JOURNAL Unpublished (1999)
COMMENT Other_GSS: BPCL-24-333H1.TJ
Contact: Shaying Zhao
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@igir.org
Clones are derived from the mouse RAC library BPCL-24. For RAC
library availability, please contact Pictor de Jong
(pdejong@mail.cho.org). Clones may be purchased from RACPAC
Resources (http://www.choi.org/bacpac/orderingframe.htm) RAC end
page: http://www.igir.org/igir\_end\_ends/mouse/24rac\_end\_infro.html
Plate: 333 row: H column: 1
Seq primer: T7
Class: RAC ends.
Location/Qualifiers
1..143
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="BPCL-24-333H1"
/clone_lib="BPCL-24"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: pTARBAC1; Site_1: BamHI; Site_2: BamHI;
BPCL-24 Mouse RAC library produced by Pictor de Jong. The
library was cloned in the pTARBAC1 cloning vector at the
BamHI sites using MboI partially digested male C57H/6J
DNA."
BASE COUNT 34 a 33 c 47 q 29 t
ORIGIN
Alignment Scores:
Pred. No.: 2,076+03 Length: 143
Score: 24.00 Matches: 5
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-09-856-070-18 (1-5) x BH101135 (1-143)

QY 1 LysGluGluLeuMet 5
DB 60 AAGGAGAGCTGATG 74

RESULT 13
LOCUS BJ008375
DEFINITION BJ008375 MF01SSA cDNA oryzias latipes cDNA clone MF01SSA19C09 5',
mRNA sequence.
ACCESSION BJ008375
VERSION BJ008375.1 GI:17368489
KEYWORDS EST.
SOURCE Japanese medaka.
ORGANISM Oryzias latipes.
REFERENCE Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Actinopterygii, Neopterygii, Teleostei, Euteleostei, Neoteleostei;
Acanthomorpha; Acanthopterygii; Perciformes; Atherinomorpha;
Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
1 (bases 1 to 149)
AUTHORS Kohara, Y., Shin-i, T., Kimura, T., Narita, T., Jindo, T. and Takeda, H.
TITLE Medaka EST Project in Takeda's lab
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasi Shin-i
Center for Genetic Resource Information
National Institute of Genetics
1111 Yatai, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
Location/Qualifiers
1..149
/organism="Oryzias latipes"
/strain="Hd-r"
/db_xref="taxon:8090"
/clone="MF01SSA19C09"
/clone_lib="MF01SSA cDNA"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="segmentation; stage 29 - 35"
BASE COUNT 43 a 33 c 38 g 33 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 2,156+03 Length: 149
Score: 24.00 Matches: 5
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-856-070-18 (1-5) x BJ008375 (1-149)

QY 1 LysGluGluLeuMet 5
DB 104 AAGGAGAGCTGATG 118

RESULT 14
LOCUS BH063341/C
DEFINITION BH063341 RIKEN full-length enriched, in vitro fertilized eggs Mus
musculus cDNA clone 7420484P13 3', mRNA sequence.
ACCESSION BH063341
VERSION BH063341.1 GI:8520754
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi;
Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae; Mus.
1 (bases 1 to 151)
AUTHORS Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci,
P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,
Hirozane, T., Horii, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,
Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,
Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kuribara, C., Kusakabe, M.,
Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,
Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata

```





Search completed: January 16, 2003, 21:37:05  
Job time : 664.571 secs

---

